Constraints on the Acquisition of Variation

Matt Goldrick & Meredith Larson
Department of Linguistics
Northwestern Institute on Complex Systems
Northwestern University
Acquiring Stochastic Patterns

• Stochastic patterns: Variability in phonological structures acquired by individual speakers.
  – Variable phonological processes
    • Word-final deletion: “left” -> /ft/ vs. /f/ (Guy 1980)
  – Variable patterns: Probabilistic phonotactics
    • Onset: /s/ more frequent than /j/ (Vitevitch et al. 2004)

• Overview of talk
  – Perspectives and evidence on the ability of speakers to acquire and represent such patterns.
  – Implicit learning experiment contrasting these perspectives.
  – Implications and future work
Probability Matching

• Claim: Any stochastic pattern that can be stated over the representational units of the system can be learned

• Implicitly assumed by many statistical theories
  – General strategy: Adjust model parameters so as to maximize the fit between model predictions and data distributions.
    – Including frameworks such as: Maximum Entropy (Goldwater & Johnson, 2002); connectionist (Dell et al., 1993); Stochastic OT (Boersma & Hayes, 2001); generalized linear models (VARBRUL; Sankoff, 1988); etc.
Problems for Probability Matching

- Niyogi: “Illogical problem of language change”
  - If learners perfectly match input, patterns should be static

- Empirical evidence: Regularization in learning
  - *Singleton & Newport (2002)*:
    - Late-learning parents exhibit variable production of ASL morphology
    - Child more closely matches systematic productions of native ASL
  - *Hudson Kam & Newport; Wonnacott & Newport (2005)*:
    - Artificial (syntactic) grammar learning task
    - Children often fail to reproduce variable input
    - Adults fail to extend variability to novel forms
An Alternative: Biases in Acquisition

- Draws on “phonetically-grounded” learning theories
  - Claim: Patterns motivated by phonetic principles are easily acquired; difficult (but not impossible!) to acquire ‘unnatural’ patterns
  - E.g., Learning to devoice final obstruents (/rad/-/>/rat/) is easier than learning a “voicing” pattern (/rat/-/>/rad/)

- Formal instantiations of phonetic grounding in acquisition
  - Substantive biases on constraint induction process (Hayes, 1999)
  - Substantive priors on Bayesian learning of constraint weights (Wilson, in press)
Biased Acquisition of Variation

• Extend this perspective to the acquisition of variation.

• Claim: Ability to represent variation is pattern-specific.
  – Certain patterns are consistent with biases, allowing learners to easily encode wide range of degrees of variation.
  – Other patterns conflict with biases, making it difficult for learners to encode stochastic patterns.
Contrasting These Perspectives

• In an implicit learning task, expose participants to data exhibiting a variety of stochastic patterns.

• Predictions
  – **Probability matching**: All stochastic patterns will be acquired.
  – **Biased acquisition**: Certain stochastic patterns will be difficult to acquire; other patterns will be easily acquired.
Implicit Learning Paradigm
(after Dell et al. 2000)

• Syllable position constraint: errors tend to preserve syllable position.
  – tap man --> map tan  > more likely than >
  – nap mat

• English: when phonotactic constraints confine a segment to syllable position, this effect is strengthened.
  – hope moan --> mope hone  > much more likely than >
  – nope moah
Implicit Learning Paradigm
(after Dell et al. 2000)

• This effect is also found when new categorical phonotactic constraints are acquired.

• Dell et al. 2000: Participants repeat tongue twisters.
  – In syllables of the tongue twisters, /f/ occurs only in onset, while /m/ occurs in both onset and coda.

• Syllable position effect found.
  – mes gek --> mes mek > more likely than >
  – mes gem

• Strengthened by categorical phonotactic constraints
  – fem ges --> fem fes > much more likely than >
  – fem gef
Current Study: Materials

• Stochastic phonotactic patterns
  – Following previous work, constraint on segment’s occurrence in onset vs. coda.

• Constrained consonant
  – /s/ or /f/, following many other studies with this paradigm

• Degree of variation

<table>
<thead>
<tr>
<th>Onset</th>
<th>20</th>
<th>40</th>
<th>50</th>
<th>60</th>
<th>80</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coda</td>
<td>80</td>
<td>60</td>
<td>50</td>
<td>40</td>
<td>20</td>
</tr>
</tbody>
</table>
Current Study: Materials

• To create CVC syllables for tongue twisters, consonant subject to stochastic phonotactic constraints combined with:

• Control consonants evenly distributed across syllable positions.
  – 6 segments, >1 feature distinct from constrained segment.
  – /k,g,m,n,p/ +
    • /v/ for /s/ conditions
    • /z/ for /f/ conditions

• Consonants subject to English phonotactics
  – /h/ and /ŋ/
Current Study: Data Collection

- Sequences of four C/ε/C syllables generated reflecting stochastic constraints of condition.
- 40 native English-speaking participants
- Sequences read aloud in time to metronome
  - Practice: 1 syllable/second
  - Error elicitation: 3 repetitions at 3.5 syllables/second
- Productions recorded and errors transcribed
  - Inter-transcriber reliability good (> 75%)
Analysis

• Question: Do participants acquire the stochastic pattern present in the data?
  – Operationalized: Is distribution of errors predicted by stochastic pattern present in the input?

• Sequence with intended coda /f/ “mef heng zek nep”
  – Error resulting in coda /f/ “zek->zef”
  – Error resulting in onset /f/ “zek->fek”

• Analysis: Are errors more likely to result in coda /f/ when /f/ is biased to appear in coda position?
  – Statistical technique: Logit regression
Input distribution is a significant predictor: F(1,18) = 10.8, p < .001
Input distribution is a significant predictor: $F(1,18) = 12.0$, $p < .001$
Input distribution is a significant predictor: $F(1,13) = 9.2,\ p < .005$
Input distribution is **NOT** a significant predictor: $F(1,15) = 1.9, \ p < .20$
The influence of input distributions on error distributions varies as a function of the stochastic pattern being acquired.
Discussion

• Participants are only selectively sensitive to stochastic patterns.
  – Problematic for probability matching perspective
  – More consistent with biased acquisition account

• Open questions
  – Nature of biases
  – How biases impact acquisition process
Frequency: No Clear Advantage for /s/-onset

- /f/ exhibits an as strong (if not stronger) preference for onset position relative to /s/.

<table>
<thead>
<tr>
<th></th>
<th>Onset</th>
<th>Coda</th>
</tr>
</thead>
<tbody>
<tr>
<td>/s/</td>
<td>52%</td>
<td>48%</td>
</tr>
<tr>
<td>/f/</td>
<td>56%</td>
<td>43%</td>
</tr>
</tbody>
</table>

Kessler & Treiman (1997):
- Type frequency, CVC monomorphemes

- Similar results in CELEX, considering
  - Token frequency
  - Multimorphemic, multisyllabic words
  - Clusters as well as singleton onset/codas
  - Only syllables with target vowel /E/

- Similar results in spontaneous speech (Buckeye corpus), taking into account possible resyllabification
Phonetic properties of /s/

• /s/ is perceptually salient
  – Greater amplitude of frication noise, more distinct spectral envelope (see Davidson, 2003, for a review)

• /s/-onset is typologically frequent
  – Special status as initial member of clusters (Morelli, 1999)

• /s/V is relatively resistant to lenition
  – /s/ weakening (s > h > ø) targets intervocalic, pre-consonantal positions (Ferguson, 1990)
Conclusions

• Speakers can acquire and represent a wide range of stochastic patterns
  – But this ability has important limitations.

• Structure of language is a synthesis of
  – The structure of environment
  – The structure of the learner

• Challenge: Understand the nature of these two forces and their interaction.
Thanks

• Transcription
  – Eva Tomczyk
  – Melissa Baese

• Running participants
  – Yaron McNabb

• SoundLab and Phonatics Group at Northwestern for advice and support